

## A MARKOV MODEL FOR THE SPREAD OF VIRUSES IN AN OPEN POPULATION

L. COUTIN,\* *Université Paul Sabatier*

L. DECREUSEFOND,\*\* *Institut Telecom*

J. S. DHERSIN,\*\* *University of Paris 13*

### Abstract

Inspired by methods of queueing theory, we propose a Markov model for the spread of viruses in an open population with an exogenous flow of infectives. We apply it to the diffusion of AIDS and hepatitis C diseases among drug users. From a mathematical point of view, the difference between the two viruses is shown in two parameters: the probability of curing the disease (which is 0 for AIDS but positive for hepatitis C) and the infection probability, which seems to be much higher for hepatitis. This model bears some resemblance to the  $M/M/\infty$  queueing system and is thus rather different from the models based on branching processes commonly used in the epidemiological literature. We carry out an asymptotic analysis (large initial population) and show that the Markov process is close to the solution of a nonlinear autonomous differential system. We prove both a law of large numbers and a functional central limit theorem to determine the speed of convergence towards the limiting system. The deterministic system itself converges, as time tends to  $\infty$ , to an equilibrium point. We then show that the sequence of stationary probabilities of the stochastic models shrinks to a Dirac measure at this point. This means that in a large population and for long-term analysis, we may replace the individual-based microscopic stochastic model with the macroscopic deterministic system without loss of precision. Moreover, we show how to compute the sensitivity of any functional of the Markov process with respect to a slight variation of any parameter of the model. This approach is applied to the spread of diseases among drug users, but could be applied to many other case studies in epidemiology.

*Keywords:* Epidemiology; Markov process; mean-field approximation

2010 Mathematics Subject Classification: Primary 60F17; 60J70; 92D30

### 1. Motivations

The hepatitis C virus (HCV for short) is spread primarily by direct contact with human blood. In developed countries that have safe blood supplies, the population infected by HCV is closely related to injecting drug users (IDUs). It is estimated that 90% of infections are due to IDUs [12]. In order to reduce the number of new hepatitis C cases, preventing infections in IDUs is then a priority. Programs exist all over the world which try to reduce the prevalence of many infectious diseases, such as HIV or hepatitis C, among IDUs. They are mainly based on needle

---

Received 15 February 2010; revision received 7 October 2010.

\* Postal address: Institut de Mathématiques de Toulouse, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France. Email address: laure.coutin@math.univ-toulouse.fr

\*\* Postal address: Telecom ParisTech, Institut Telecom, 46 rue Barrault, 75634 Paris Cedex 13, France.

Email address: laurent.decreusefond@telecom-paristech.fr

\*\*\* Postal address: Department of Mathematics, Institute Galilée, University of Paris 13, 99 Avenue Jean-Baptiste Clément, F-93430 Villetaneuse, France. Email address: dherstin@math.univ-paris13.fr

exchanges. It turns out that after several years of such programs, the HIV prevalence seems to be now rather low, whereas the percentage of IDUs who are HCV positive remains about 60% [12], [15]. We were asked by epidemiologists to provide a mathematical model which could quantitatively evaluate the differences between the two diseases. From a medical point of view, there are two main differences between HIV and HCV. On the one hand, we can cure hepatitis C patients to the point where the patient becomes seronegative; on the other hand, HCV resists air contact much more than HIV and, thus, is more contagious during needle sharing. The actual question raised by epidemiologists was whether this can be the only explanation for the high prevalence of the HCV or whether there exists another parameter which could be more easily controlled to reduce this prevalence.

It is always a challenge to analyze an epidemic problem because there are many real-life situations that should be incorporated while keeping the mathematical model tractable. Moreover, epidemic field studies are expensive and hard to organize, so parameter estimates are rare and often imprecise. It is thus necessary to deal with parsimonious models whose parameters have clear and visible meanings. To the best of the authors' knowledge, the only models which have been developed for the dynamics of HCV transmission are found in [8] and [19]. It is a deterministic model with more than twenty-five parameters, for which the authors do not have explicit results for the asymptotics and only estimate them by simulations.

In our paper, we thus propose a Markovian model for the spread of HCV/HIV in an open local population of IDUs. That is, we consider one household (see [1] and the references therein for epidemiological vocabulary and classical models); the effect of other households is summarized by a parameter  $r$  which represents the incidence of infectives coming from the outside, i.e. other households. Our model differs from the classical SIR, SIRS, SIS, etc. models in several aspects (see [2], [5], [6], and [13] for the most similar models). There are two 'degrees' of susceptibility. The first category of susceptibility are the people who can potentially become IDUs, i.e. everybody. A second category is composed of IDUs not yet infected. Hence, the system is open: there exists a continuous influx of new IDUs. Another difference with usual models is that some IDUs arrive in the studied population as infected and, thus, do not pass through the usual state of susceptibility. Moreover, there is no recovered category in our model since we cannot measure their number (when they are no longer an IDU, they cannot be counted in studies focused on drug users), but some infected users can recover (from the HCV) and return to the state of susceptible IDUs. Some epidemic models use branching processes. It seems rather hard to proceed along this line here, notably because the rate of infection changes with the situation of the whole population and is not determined locally for each leaf. Here, we have an approach inspired by queueing theory. Specifically, our model bears some strong resemblance to the  $M/M/\infty$  queue. As is usual for such processes, with the Markov model being intractable in the sense that we cannot compute its stationary distribution, we resort to a large population analysis. We prove that in such a situation, the prevalence can be well represented by a deterministic differential system. Borrowing ideas from statistical physics, this means that, for a large population, the macroscopic dynamic of the disease can be well approximated by that of a deterministic system. Furthermore, it turns out that the deterministic system has a unique limiting point. We then devote some time to prove that the stationary distribution of the finite-population Markov processes converges to a Dirac measure located at this fixed point. This says that, as far as the long-time behavior is concerned, the statistics of the Markov process are close to that of the fixed point of the deterministic system.

To keep the Markovian character of our model, we make the following usual and reasonable hypothesis. Exogenous antibody-positive and antibody-negative drug users arrive in this local population according to Poisson processes. If initiated by an antibody-positive drug addict, a new IDU acquires the virus very rapidly after the initiation [3], [10]. HCV or HIV then spreads in the population through the sharing of syringes, needles, and other accessories (cotton, boilers, etc.). Each individual of the population stays in his state (infected/noninfected) for an exponentially distributed time. We present the model in Section 3. The difference between HIV and HCV diffusions lies only in the parameters: the probability of curing the disease (which is 0 for AIDS but positive for hepatitis C) and the infection probability, which seems to be much higher for hepatitis. If we respectively denote by  $X_1(t)$  and  $X_2(t)$  the numbers of antibody-positive and antibody-negative individuals in the local population at time  $t$ , we prove that the process  $X = (X_1, X_2)$  is an ergodic Markov process. In Section 4 we give a related deterministic differential system connected with this Markov process. We study its asymptotic behavior and give an explicit expression of the limit of the solution. In Section 5 we give a mean-field approximation of the process  $X$ : for large populations, we prove that the process  $X$  is close to the solution  $\psi$  of the deterministic differential system. In Section 6 we prove that, for large populations, the invariant distribution for the Markov process  $X$  can be approximated by the Dirac measure which only charges  $\psi(\infty)$ . Hence, we can give an explicit limit of the prevalence of a virus in the population. In Section 7 we give a central limit theorem for the approximation of  $X$  by  $\psi$  when the population tends to  $\infty$ . In Section 8 we show that even for relatively small values (around one hundred) of the initial population, there is a good accordance between the prevalence computed on the deterministic limit and the prevalence observed in the stochastic model. We also show that this can be extended to the sensitivity of the model with respect to slight variations of some parameters.

### 2. Preliminaries

Let us denote by  $\mathbb{D}([0, T], \mathbb{R}^2)$  the set of càdlàg functions (those that are continuous from the right with left limits) equipped with its usual topology. In this section we recall some results about càdlàg semimartingales; for details, we refer the reader to [16]. We assume that we are given  $(\Omega, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ , a filtered probability space satisfying the so-called usual hypothesis. On  $(\Omega, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ , let  $X$  and  $Y$  be two real-valued càdlàg square integrable semimartingales. The mutual variation of  $X$  and  $Y$ , denoted by  $[X, Y]$ , is the right-continuous process with finite variation such that the following integration by parts formula is satisfied:

$$X(t)Y(t) - X(0)Y(0) = \int_{(0,t]} X(s_-) dY(s) + \int_{(0,t]} Y(s_-) dX(s) + [X, Y]_t.$$

The Meyer process of the couple  $(X, Y)$ , or its square bracket, denoted by  $\langle X, Y \rangle$ , is the unique right-continuous with finite variation predictable process such that

$$X(t)Y(t) - X(0)Y(0) - \langle X, Y \rangle_t$$

is a martingale. Alternatively,  $\langle X, Y \rangle$  is the unique right-continuous, predictable process with finite variation such that  $[X, Y] - \langle X, Y \rangle$  is a martingale. For a vector-valued semimartingale  $X = (X_1, X_2)$ , where  $X_1$  and  $X_2$  are real-valued martingales, we denote by  $\langle\langle X \rangle\rangle$ , its square bracket, defined by

$$\langle\langle X \rangle\rangle_t = \begin{pmatrix} \langle X_1 \rangle_t & \langle X_1, X_2 \rangle_t \\ \langle X_1, X_2 \rangle_t & \langle X_2 \rangle_t \end{pmatrix}.$$

In the sequel, if  $x$  is a vector or  $M$  is a matrix, we denote by  $\|x\|$  or, respectively,  $\|M\|$  its  $L^1$ -norm.

Let  $E$  be a discrete denumerable space. Let  $(X(t), t \geq 0)$  be an  $E$ -valued, pure-jump Markov process, with infinitesimal generator  $Q = (q_{xy}, (x, y) \in E \times E)$ . For any  $F: E \rightarrow \mathbb{R}$ , Dynkin's lemma states that the process

$$F(X(t)) - F(X(0)) - \int_0^t QF(X_s) ds$$

is a local martingale, where

$$QF(x) = \sum_{y \neq x} (F(y) - F(x))q_{xy}.$$

Here and hereafter, we identify the matrix  $Q$  and the operator  $Q$  defined as above.

### 3. Markov model

We consider the dynamics of HCV/HIV among a local open population. There is a constant rate  $r$  of immigration of antibody-positive drug users from other communities. These arrivals are modeled as a Poisson process of intensity  $r$ . We let  $X_1(t)$  and  $X_2(t)$  respectively denote the numbers of antibody-positive and antibody-negative users at time  $t$  in the population under consideration. The new susceptible drug users arrive as a Poisson process of intensity  $\lambda$ . We assume that, for their first injection, they are initiated by an older IDU who has a probability  $q(t) = X_1(t)(X_1(t) + X_2(t))^{-1}$  of being infected. For different reasons, even in this situation, the probability of being infected is not exactly 1 and is denoted by  $p_1$ . Each time an antibody-negative IDU injects he/she may share some of his paraphernalia and may become infected if the sharing occurs with an infected IDU. We summarize all these probabilities by saying that at each injection, the probability of becoming infected is  $pq(t)$ , where  $p$  is a parameter to be estimated, as is  $p_1$ . If we denote by  $\alpha$  the rate at which an IDU injects drugs, and if  $\alpha p$  is small, we can assume that the rate at which a sane IDU in the population is infected is given by  $\alpha pq(t)$ . Once infected, an IDU may either die, stop drug usage, or self heal. The time for this to happen is modeled by an exponential distribution of parameter  $\mu_1$ . There is a probability  $\gamma$  that the first two events happen and, thus, a probability of  $1 - \gamma$  that he/she self heals and then appears in the system as a new seronegative. For an antibody-negative IDU, the only way to exit the population is by stopping injecting drugs or dying, which is supposed to happen after an exponentially distributed duration with parameter  $\mu_2$ . In summary, the transitions are described in Figure 1.

**Remark 3.1.** Note that the main difference between the hepatitis C and AIDS viruses lies in the values of  $p$  and  $p_1$  (which is supposed to be much less for the AIDS virus) and in the value of  $\gamma$ , which is a priori 0 for the AIDS virus.

In what follows, we set

$$q_1(n_1, n_2) = r + \lambda p_1 \frac{n_1}{n_1 + n_2}, \quad q_2(n_1, n_2) = \gamma \mu_1 n_1, \quad q_3(n_1, n_2) = \alpha p n_2 \frac{n_1}{n_1 + n_2},$$

$$q_4(n_1, n_2) = \lambda \left( 1 - p_1 \frac{n_1}{n_1 + n_2} \right), \quad q_5(n_1, n_2) = \mu_2 n_2, \quad q_6(n_1, n_2) = (1 - \gamma) \mu_1 n_1.$$

Note that here and hereafter, we use the convention  $0/0 = 0$ .

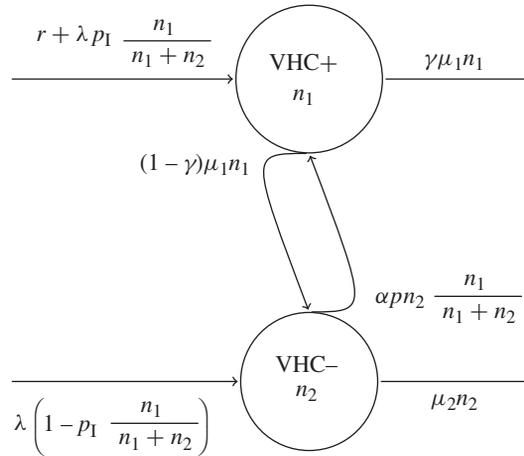


FIGURE 1: Transitions of the Markov process from and to state  $(n_1, n_2)$ , where VHC+ and VHC- respectively refer to infected and uninfected viral hepatitis C users.

**Lemma 3.1.** Let  $x^0 = (x_1^0, x_2^0)$ . Conditionally on  $X(0) = x^0$ , the process

$$W(t) = X_1(t) + X_2(t) - (x_1^0 + x_2^0)$$

is dominated (for the strong stochastic order of processes) by a Poisson process of intensity  $r + \lambda$ . In particular, for any  $p \geq 1$ , there exists a universal constant  $C_p$  such that, for any  $T > 0$ ,

$$\mathbb{E} \left[ \sup_{t \leq T} \|X(t)\|^p \mid X(0) = x^0 \right] \leq C_p (\|x^0\|^p + (r + \lambda)^p T^p).$$

*Proof.* It suffices to say that by suppressing all the departures we obtain another system with a population larger than that of the system under consideration, at any time, for any trajectory. Then,  $X_1(t) + X_2(t) - (x_1^0 + x_2^0)$  is less than the number of arrivals of a Poisson process of intensity  $r + \lambda$ . Since a Poisson process has increasing paths, its supremum over  $[0, T]$  is its value at time  $T$ . The second assertion follows.

**Theorem 3.1.** The Markov process  $X = (X_1, X_2)$  is ergodic. For  $r > 0$ , the process  $X$  is irreducible. For  $r = 0$ , the set  $\{(n_1, n_2) \in \mathbb{N} \times \mathbb{N}, n_1 = 0\}$  is a proper closed subset.

*Proof.* Let  $S$  be the function defined on  $\mathbb{N} \times \mathbb{N}$  by

$$S(n_1, n_2) = \|(n_1, n_2)\| = n_1 + n_2.$$

If we denote by  $Q$  the infinitesimal generator of  $X$ , we have

$$QS(n_1, n_2) = \lambda + r - \gamma \mu_1 n_1 - \mu_2 n_2.$$

Let  $K$  be real and strictly greater than  $(\lambda + r + 1) / \mu_-$ , where  $\mu_- = \min(\gamma \mu_1, \mu_2)$ , and consider the following finite subset of the state space:

$$D_K = \{(n_1, n_2) \in \mathbb{N} \times \mathbb{N}, n_1 + n_2 \leq K\}.$$

If  $(n_1, n_2)$  belongs to  $D_K^c$  then

$$QS(n_1, n_2) \leq \lambda + r - \mu_-(n_1 + n_2) < -1.$$

Lemma 3.1 implies that both

$$E \left[ \sup_{s \in [0,1]} S(X(s)) \right] \quad \text{and} \quad E \left[ \int_0^1 |QS(X(s))| ds \right]$$

are finite. Then, according to [17, Proposition 8.14],  $X$  is ergodic.

The second and third assertions are immediate through inspection of the transition rates.

With the nonlinearity appearing in the transitions, it seems hopeless to find an exact expression for the stationary probability of the Markov process  $(X_1, X_2)$ . As is usual in queueing theory [17], we then resort to an asymptotic analysis in order to gain some insights into the evolution of this system. This means that we let the initial population grow larger and larger. Keeping other quantities of the same order of magnitude, we are thus led to increase  $r$  and  $\lambda$  at the same speed, i.e. keeping the ratio  $i = (r + \lambda)/(x_1^0 + x_2^0)$  constant. Note that in the epidemiological terminology,  $i$  is the incidence of a new susceptible. It is measured in the percentage of individuals per year.

#### 4. A deterministic differential system

The mean-field approximation will lead us to investigate the solutions of the following differential system with initial condition  $x^0 = (x_1^0, x_2^0) \in (\mathbb{R}_+ \times \mathbb{R}_+) \setminus \{(0, 0)\}$ :

$$\begin{aligned} \psi_1'(t) &= r + \lambda p_I \frac{\psi_1(t)}{\psi_1(t) + \psi_2(t)} - \mu_1 \psi_1(t) + \alpha p \frac{\psi_1(t) \psi_2(t)}{\psi_1(t) + \psi_2(t)}, \\ \psi_1(0) &= x_1^0, \\ \psi_2'(t) &= \lambda \left( 1 - p_I \frac{\psi_1(t)}{\psi_1(t) + \psi_2(t)} \right) - \mu_2 \psi_2(t) - \alpha p \frac{\psi_1(t) \psi_2(t)}{\psi_1(t) + \psi_2(t)} + (1 - \gamma) \mu_1 \psi_1(t), \\ \psi_2(0) &= x_2^0. \end{aligned}$$

We denote this system by  $S_r(x^0)$ .

**Theorem 4.1.** *For any  $x^0 = (x_1^0, x_2^0) \in (\mathbb{R}_+ \times \mathbb{R}_+) \setminus \{(0, 0)\}$ , there exists a unique solution to  $S_r(x^0)$ . Furthermore, this solution is defined for any time. For  $r > 0$ , the differential system has a unique fixed point  $(\xi_1, \xi_2)$  in  $\mathbb{R}_+ \times \mathbb{R}_+$ , defined by*

$$\xi_2 = \frac{1}{\mu_2} (r + \lambda - \gamma \mu_1 \xi_1), \quad 0 = \mu_1 (\gamma (\alpha p - \mu_1) + \mu_2) \xi_1^2 - b \xi_1 + r (\lambda + r), \quad \text{and} \quad \xi_1 \geq 0, \tag{4.1}$$

where

$$b = \lambda (\alpha p - \mu_1 + p_I \mu_2) + r (\mu_2 - (1 + \gamma) \mu_1 + \alpha p).$$

Moreover, for  $r > 0$  and any  $x^0 \in \mathbb{R}_+^2 \setminus \{0, 0\}$ ,

$$\lim_{t \rightarrow +\infty} (\psi_1(t), \psi_2(t)) = (\xi_1, \xi_2).$$

*Proof.* We denote by  $f_1$  and  $f_2$  the functions that allow  $S_r(x^0)$  to be written as

$$\psi'_1(t) = f_1(\psi_1(t), \psi_2(t)) \quad \text{and} \quad \psi'_2(t) = f_2(\psi_1(t), \psi_2(t)).$$

Since  $f_1$  and  $f_2$  are locally Lipschitz, there exists a local solution for any starting point  $x^0$  belonging to  $(\mathbb{R}_+ \times \mathbb{R}_+) \setminus \{(0, 0)\}$ . Moreover, for any  $(x_1, x_2) \in (\mathbb{R}_+ \times \mathbb{R}_+) \setminus \{(0, 0)\}$ ,

$$r - \mu_1 x_1 \leq f_1(x_1, x_2) \leq r + \lambda p_1 + \alpha p x_1,$$

$$\text{and} \quad \lambda(1 - p_1) - \mu_2 x_2 \leq f_2(x_1, x_2) \leq \lambda + (1 - \gamma)\mu_1 x_1.$$

By standard theorems about the comparison of solutions of differential equations, we can show that every local solution  $\psi$  can be extended to  $\mathbb{R}$  and that, for any  $t \in \mathbb{R}$ ,  $\psi(t) = (\psi_1(t), \psi_2(t))$  belongs to  $(\mathbb{R}_+ \times \mathbb{R}_+) \setminus \{(0, 0)\}$ . Furthermore, with direct calculations, we have

$$\frac{d}{dt}(\psi_1(t) + \psi_2(t)) = r + \lambda - \gamma\mu_1\psi_1(t) - \mu_2\psi_2(t). \tag{4.2}$$

For  $\varepsilon > 0$ , consider

$$A^\pm_\varepsilon = \{(x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}_+, 0 \leq \pm(r + \lambda - \gamma\mu_1 x_1 - \mu_2 x_2) < \varepsilon\},$$

$$B^\varepsilon_+ = \{(x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}_+, r + \lambda - \gamma\mu_1 x_1 - \mu_2 x_2 \geq \varepsilon\},$$

$$B^\varepsilon_- = \{(x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}_+, r + \lambda - \gamma\mu_1 x_1 - \mu_2 x_2 \leq -\varepsilon\},$$

and

$$A^0 = \{(x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}_+, r + \lambda - \gamma\mu_1 x_1 - \mu_2 x_2 = 0\}.$$

According to (4.2), on  $B^\varepsilon_+$ , the derivative of  $\psi_1 + \psi_2 = \|(\psi_1, \psi_2)\|$  is greater than  $\varepsilon$ ; hence, for a starting point in  $A^\varepsilon_+$ , the trajectory has an  $L^1$  increasing norm. Reasoning along the same lines on  $B^\varepsilon_-$ , we see that, for any  $\eta > 0$ , for any starting point outside  $A^0$ , the trajectory of the differential system enters, in a finite time, the set  $A^\eta_+$  or  $A^\eta_-$ . Moreover, after this time, the orbit stays in the compact  $A^\eta_+ \cup A^\eta_-$  forever. It follows that (see, for instance, [18])

$$\lim_{t \rightarrow +\infty} \text{dist}((\psi_1(t), \psi_2(t)), A^0) = 0.$$

This implies that any invariant set  $M$  must be included in  $A^0$ . We then seek a maximal invariant set. It is given by the intersection of the sets  $Z_i = \{(x_1, x_2), f_i(x_1, x_2) = 0\}$ ,  $i = 1, 2$ . We then remark that this system of equations is equivalent to the system  $f_1 + f_2 = 0$  and  $f_1 = 0$ . It turns out that

$$(f_1 + f_2)(x_1, x_2) = r + \lambda - \gamma\mu_1 x_1 - \mu_2 x_2 = 0.$$

The equation  $f_1(x_1, x_2) = 0$  yields

$$x_2 = \frac{\mu_1 x_1(x_1 - (\lambda p_1 + r)/\mu_1)}{r + (\alpha p - \mu_1)x_1} = h(x_1).$$

Assume first that  $r > 0$  if  $\alpha p - \mu_1 \geq 0$ . Then  $h$  is negative for  $x_1 \in [0, c)$  and  $c = (\lambda p_1 + r)/\mu_1$ , and  $h$  is strictly increasing and nonnegative for  $x_1 \geq c$ . Then there exists one and only one equilibrium point whose coordinates  $(\xi_1, \xi_2)$  are thus given by the solution of (4.1)—see Figure 2 for an illustration.

If  $\alpha p - \mu_1 < 0$  then  $h$  has an infinite branch at  $z^1 = -r/(\alpha p - \mu_1)$ . Then  $z^2 = (r + \lambda p_1)/\mu_1 \leq z^1$  and  $h$  is negative before  $z^2$  and after  $z^1$ , and  $h$  is a strictly increasing

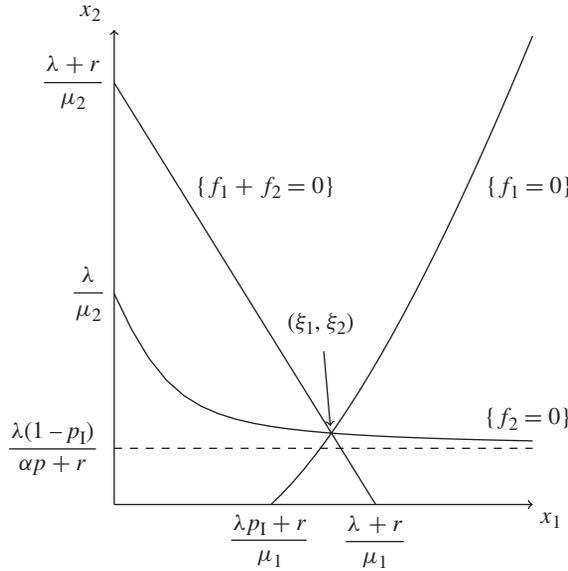


FIGURE 2: Determination of the fixed point.

homeomorphism from  $[z^2, z^1)$  onto  $\mathbb{R}^+$ ; thus, there exists one and only one solution to (4.1). Alternatively,  $z^2 > z^1$  and so  $h$  is positive only between  $z^1$  and  $z^2$ ; thus, again, there exists a unique fixed point.

Since  $A^0$  is a line, it is invariant if and only if at any time the tangent vector is null or colinear to  $(\mu_2, -\gamma\mu_1)$ . It is trivially true for  $(x_1^0, x_2^0) = (\xi_1, \xi_2)$ . For any  $(x_1^0, x_2^0) \in A^0 \setminus \{(\xi_1, \xi_2)\}$ , in view of (4.2),  $(\psi_1'(0), \psi_2'(0))$  is colinear to  $(1, -1)$ .

We then distinguish two situations: either  $\gamma\mu_1 = \mu_2$  or not. If  $\gamma\mu_1 \neq \mu_2$ , we see that  $A^0 \setminus \{(\xi_1, \xi_2)\}$  is not invariant and, thus, according to the Poincaré–Bendixson theorem,  $M = \{(\xi_1, \xi_2)\}$ . If  $\gamma\mu_1 = \mu_2$  then  $A^0$  is invariant. However,  $\psi_1 + \psi_2$  is a solution of the differential equation

$$v'(t) = r + \lambda - 2\gamma\mu_1 v(t), \quad v(0) = x_1^0 + x_2^0.$$

By direct integration, this yields

$$(\psi_1 + \psi_2)(t) = (x_1^0 + x_2^0)e^{-2\gamma\mu_1 t} + \frac{r + \lambda}{2\gamma\mu_1}(1 - e^{-2\gamma\mu_1 t}).$$

Since  $A^0$  is compact, there exists a minimum invariant set, say  $M$ . According to the Poincaré–Bendixson theorem,  $M$  is either a periodic orbit or a critical point. Since  $\psi_1 + \psi_2$  is not periodic,  $M$  is also reduced to  $(\xi_1, \xi_2)$ . This completes the proof.

For the sake of completeness, we also describe what happens for  $r = 0$ . This corresponds to an IDU population where only sane new IDUs arrive. The proof is similar to that of Theorem 4.1.

**Theorem 4.2.** *With the same notation as above, if  $r = 0$  and  $x_1^0 = 0$ , then*

$$\psi_1(t) = 0 \text{ for all } t, \quad \lim_{t \rightarrow +\infty} (\psi_1(t), \psi_2(t)) = \left(0, \frac{\lambda}{\mu_2}\right).$$

If  $r = 0$  and  $\rho = -\gamma(\alpha p - \mu_1) - \mu_2 > 0$ , then there exists two equilibrium points:  $(0, \lambda/\mu_2)$  and

$$\xi_1 = \frac{\lambda \alpha p - \mu_1 + \mu_2 p_1}{\rho}, \quad \xi_2 = \frac{\lambda}{\rho}(1 + \gamma p_1).$$

If  $x_1^0 > 0$  then

$$\lim_{t \rightarrow +\infty} (\psi_1(t), \psi_2(t)) = (\xi_1, \xi_2).$$

If  $r = 0$  and  $\rho > 0$ , then, for any  $x^0$  with positive  $x_1^0$ ,

$$\lim_{t \rightarrow +\infty} (\psi_1(t), \psi_2(t)) = \left(0, \frac{\lambda}{\mu_2}\right).$$

In what follows, we denote by  $\psi^\infty$  the unique point to which the system converges in each case. We denote by  $\Psi$  the measurable function such that  $\Psi(x^0, t)$  is the value of the solution to  $S_r(x^0)$  at time  $t$ .

### 5. Mean-field approximation

We now consider a sequence  $(X^N(t) = (X_1^N(t), X_2^N(t)), t \geq 0)$  of Markov processes with the same transitions as above but with different rates given by (with self-evident notation):

$$\begin{aligned} q_1^N(n_1, n_2) &= r_N + \lambda_N p_1 \frac{n_1}{n_1 + n_2}, \\ q_2^N(n_1, n_2) &= \gamma \mu_1 n_1, \\ q_3^N(n_1, n_2) &= \alpha p n_2 \frac{n_1}{n_1 + n_2}, \\ q_4^N(n_1, n_2) &= \lambda_N \left(1 - p_1 \frac{n_1}{n_1 + n_2}\right), \\ q_5^N(n_1, n_2) &= \mu_2 n_2, \\ q_6^N(n_1, n_2) &= (1 - \gamma) \mu_1 n_1. \end{aligned}$$

The main result of this section is the following mean-field approximation of the system  $X^N$ .

**Theorem 5.1.** Assume that, as  $N \rightarrow +\infty$ ,

$$\mathbb{E} \left[ \left\| \frac{1}{N} X^N(0) - x^0 \right\|^2 \right] \rightarrow 0, \quad \frac{1}{N} r_N \rightarrow r \geq 0, \quad \frac{1}{N} \lambda_N \rightarrow \lambda.$$

Let  $\Psi(x^0, \cdot) = (\Psi_1(x^0, \cdot), \Psi_2(x^0, \cdot))$  be the solution of the differential system  $S_r(x^0)$ . Then, for any  $T > 0$ ,

$$\mathbb{E} \left[ \sup_{t \leq T} \left\| \frac{1}{N} X^N(t) - \Psi(x^0, t) \right\|^2 \right] \rightarrow 0 \quad \text{as } N \rightarrow +\infty.$$

Before proving Theorem 5.1, we give the martingale problem satisfied by the process  $X^N$ .

**Theorem 5.2.** For any  $N > 0$ , the process  $X^N$  is a vector-valued semimartingale with decomposition

$$\begin{aligned} X_1^N(t) &= X_1^N(0) + \int_0^t (q_1^N + q_3^N - q_2^N - q_6^N)(X^N(s)) \, ds + M_1^N(t), \\ X_2^N(t) &= X_2^N(0) + \int_0^t (q_4^N - q_3^N - q_5^N + q_6^N)(X^N(s)) \, ds + M_2^N(t), \end{aligned}$$

where  $M^N = (M_1^N, M_2^N)$  is a local martingale vanishing at 0 with square bracket given by

$$\langle\langle M^N \rangle\rangle_t = \begin{pmatrix} \int_0^t (q_1^N + q_3^N + q_2^N + q_6^N)(X^N(s)) \, ds & - \int_0^t (q_3^N + q_6^N)(X^N(s)) \, ds \\ - \int_0^t (q_3^N + q_6^N)(X^N(s)) \, ds & \int_0^t (q_4^N + q_3^N + q_5^N + q_6^N)(X^N(s)) \, ds \end{pmatrix}.$$

*Proof.* Using the martingale problem associated with the Markov process  $X^N$ , we obtain, for  $t \geq 0$ ,

$$X^N(t) = X^N(0) + \begin{pmatrix} \int_0^t (q_1^N + q_3^N - q_2^N - q_6^N)(X^N(s)) \, ds \\ \int_0^t (q_4^N - q_3^N - q_5^N + q_6^N)(X^N(s)) \, ds \end{pmatrix} + M^N(t),$$

where  $M^N = (M_1^N, M_2^N)$  is a two-dimensional local martingale vanishing at 0.

We now compute its square bracket. First of all, we consider  $\langle M_1^N, M_2^N \rangle$ . Using integration by parts, we obtain, for  $t \geq 0$ ,

$$X_1^N(t)X_2^N(t) = X_1^N(0)X_2^N(0) + \int_{(0,t]} X_1^N(s_-) \, dX_2^N(s) + \int_{(0,t]} X_2^N(s_-) \, dX_1^N(s) + [X_1^N, X_2^N]_t,$$

where  $[X_1^N, X_2^N]$  denotes the mutual variation of  $X_1^N$  and  $X_2^N$ . Hence,

$$\begin{aligned} X_1^N(t)X_2^N(t) &= X_1^N(0)X_2^N(0) + \int_0^t X_1^N(s)(q_4^N - q_3^N - q_5^N + q_6^N)(X^N(s)) \, ds \\ &\quad + \int_0^t X_2^N(s)(q_1^N + q_3^N - q_2^N - q_6^N)(X^N(s)) \, ds + [X_1^N, X_2^N]_t \\ &\quad + \text{local martingale.} \end{aligned}$$

Now, writing the martingale problem associated with the process  $X_1^N X_2^N$ , we have

$$\begin{aligned} X_1^N(t)X_2^N(t) &= X_1^N(0)X_2^N(0) + \int_0^t X_1^N(s)(q_4^N - q_5^N)(X^N(s)) \, ds \\ &\quad + \int_0^t X_2^N(s)(q_1^N - q_2^N)(X^N(s)) \, ds \\ &\quad + \int_0^t (X_2^N(s) - X_1^N(s) - 1)q_3^N(X^N(s)) \, ds \\ &\quad + \int_0^t (X_1^N(s) - X_2^N(s) - 1)q_6^N(X^N(s)) \, ds + \text{local martingale.} \end{aligned}$$

We conclude that

$$\langle X_1^N, X_2^N \rangle_t = - \int_0^t (q_3^N + q_6^N)(X^N(s)) \, ds.$$

Similar arguments show that

$$\langle X_1^N \rangle_t = \int_0^t (q_1^N + q_3^N + q_2^N + q_6^N)(X^N(s)) \, ds,$$

$$\text{and } \langle X_2^N \rangle_t = \int_0^t (q_4^N + q_3^N + q_5^N + q_6^N)(X^N(s)) \, ds,$$

which completes the proof.

*Proof of Theorem 5.1.* According to Theorem 4.1, for any  $x^0 \in \mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}$ ,  $\inf_{s \in \mathbb{R}_+} \|\Psi(x^0, s)\| > 0$ . Then, Theorem 5.1 is a consequence of the following lemma.

**Lemma 5.1.** *There exists a constant  $C$  depending only on  $r, \lambda, p_1, \mu_1, \mu_2$ , and  $\alpha p$  such that, for any  $x^0 \in \mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}$ , any  $(X^M(0))_{M \in \mathbb{N}}$  sequence of random variables taking values in  $\mathbb{R} \times \mathbb{R} \setminus \{(0, 0)\}$ , any  $N \in \mathbb{N}^*$ , and any  $T > 0$ ,*

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} \left\| \frac{1}{N} X^N(t) - \Psi(x^0, t) \right\|^2 \mid \sigma(X^M(0), M \in \mathbb{N}) \right] \\ & \leq \left( \left\| \frac{1}{N} X^N(0) - x^0 \right\|^2 + \frac{1}{N} \left( T + T^2 \frac{1}{N} \|X^N(0)\| \right) \right) \\ & \quad \times \exp \left( T \int_0^T \left( 1 + \frac{1}{\|\Psi(x^0, s)\|} \right)^2 \, ds \right). \end{aligned}$$

*Proof.* Let us fix  $T > 0$ . Using Theorem 5.2, we have

$$\begin{aligned} \frac{1}{N} X_1^N(t) &= \frac{1}{N} X_1^N(0) + \int_0^t \frac{1}{N} (q_1^N + q_3^N - q_2^N - q_6^N)(X^N(s)) \, ds + \frac{1}{N} M_1^N(t), \\ \frac{1}{N} X_2^N(t) &= \frac{1}{N} X_2^N(0) + \int_0^t \frac{1}{N} (q_4^N - q_3^N - q_5^N + q_6^N)(X^N(s)) \, ds + \frac{1}{N} M_2^N(t). \end{aligned}$$

Moreover,

$$\psi_1(t) = \int_0^t (q_1 + q_3 - q_2 - q_6)(\psi(s)) \, ds, \quad \psi_2(t) = \int_0^t (q_4 - q_3 - q_5 + q_6)(\psi(s)) \, ds.$$

Note that, for  $x = (x_1, x_2)$  in  $\mathbb{R}_+ \times \mathbb{R}_+$  and  $y = (y_1, y_2)$  in  $\mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}$ ,

$$\begin{aligned} \left| \frac{x_1}{x_1 + x_2} - \frac{y_1}{y_1 + y_2} \right| &\leq \left| \frac{x_1 - y_1}{y_1 + y_2} \right| + \left| \frac{x_1}{x_1 + x_2} - \frac{y_1}{y_1 + y_2} \right| \\ &= \left| \frac{x_1 - y_1}{y_1 + y_2} \right| + \left| \frac{x_1}{x_1 + x_2} \frac{y_1 - x_1 + y_2 - x_2}{y_1 + y_2} \right| \\ &\leq 2 \frac{\|x - y\|}{\|y\|}. \end{aligned}$$

We also have

$$\left| \frac{x_1 x_2}{x_1 + x_2} - \frac{y_1 y_2}{y_1 + y_2} \right| \leq 2 \|x - y\|.$$

From now on, we use  $C$  to denote a positive constant which depends only on  $r, \lambda, p_1, \mu_1, \gamma, \mu_2$ , and  $\alpha p$ , and which may vary from line to line. For  $0 \leq t \leq T$ ,

$$\begin{aligned} & \left\| \frac{1}{N} X^N(t) - \Psi(x^0, t) \right\|^2 \\ & \leq C \left( \left\| \frac{1}{N} X^N(0) - \Psi(x^0, 0) \right\|^2 + T^2 \left| r - \frac{r_N}{N} \right|^2 + T^2 \left| \lambda - \frac{\lambda_N}{N} \right|^2 \right. \\ & \quad \left. + T \int_0^t \left( 1 + \frac{1}{\|\Psi(x^0, s)\|} \right)^2 \left\| \frac{1}{N} X^N(s) - \Psi(x^0, s) \right\|^2 ds + \frac{1}{N^2} \left\| M^N(t) \right\|^2 \right). \end{aligned} \tag{5.1}$$

Using the Burkholder–Davis–Gundy inequality [16], we obtain

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \|M^N(t)\|^2 \mid \sigma(X^M(0), M \in \mathbb{N}) \right] \leq C \mathbb{E}[\|M^N\|_T \mid \sigma(X^M(0), M \in \mathbb{N})],$$

where, for a square matrix  $A$ ,  $|A| = |\text{trace}(A)|$ . As a consequence of Lemma 3.1, we obtain, for  $i \in \{1, 2, 3, 4, 5\}$ ,

$$\mathbb{E} \left[ \sup_{t \leq T} q_i^N(X_s) \right] \leq C(\|X^N(0)\| + (r_N + \lambda_N)T) \tag{5.2}$$

and

$$\mathbb{E}[\|M^N\|_T \mid \sigma(X^M(0), M \in \mathbb{N})] \leq CT(\mathbb{E}[\|X^N(0)\|] + (r_N + \lambda_N)T).$$

Hence, using Gronwall’s lemma, (5.1) implies that

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} \left\| \frac{1}{N} X^N(t) - \Psi(x^0, t) \right\|^2 \mid \sigma(X^M(0), M \in \mathbb{N}) \right] \\ & \leq \left( \left\| \frac{1}{N} X^N(0) - x^0 \right\|^2 + \frac{1}{N} \left( T + T^2 \frac{1}{N} \|X^N(0)\| \right) \right) \\ & \quad \times \exp \left( T \int_0^T \left( 1 + \frac{1}{\|\Psi(x^0, s)\|} \right)^2 ds \right). \end{aligned}$$

This completes the proof.

### 6. Stationary regime

We have proved so far that the process  $N^{-1} X^N$  converges, as  $N$  tends to  $\infty$ , to a deterministic  $\mathbb{R}^2$ -valued function. This function converges, as  $t$  tends to  $\infty$ , to a fixed point  $\psi^\infty$ . On the other hand, for each  $N$ , the Markov process  $X^N$  is ergodic and, thus, has a limiting distribution as  $t$  tends to  $\infty$ . This raises the natural question of whether this limiting distribution converges to the Dirac mass at  $\psi^\infty$  as  $N$  tends to  $\infty$ . Let us denote by  $P_{Y^N, \nu}$  the distribution of the process  $Y^N = N^{-1} X^N$  under initial distribution  $\nu$ . We denote by  $P_{\psi, \nu}$  the distribution of the process whose initial state is chosen according to  $\nu$  and whose deterministic evolution is then given by the differential system  $S_r(x^0)$ . According to Theorem 3.1, we know that  $X^N$  has a stationary probability whose value is irrespective of the initial distribution of  $X^N$ . We denote by  $Y^N(\infty)$

a random variable whose distribution is the stationary measure of  $Y^N$ . What we already know can be summarized in the following diagram.

$$\begin{array}{ccc}
 P_{Y^N(t), \delta_{x^0}} & \xrightarrow{N \rightarrow \infty} & P_{\Psi(\cdot, t), \delta_{x^0}} \\
 t \rightarrow \infty \downarrow & & \downarrow t \rightarrow \infty \\
 P_{Y^N(\infty)} & \xrightarrow[?]{N \rightarrow \infty} & \delta_{\psi^\infty}
 \end{array}$$

The problem is then to prove that this is a commutative diagram, i.e. that  $Y^N(\infty)$  converges in distribution to the equilibrium point of the system  $S_r(x^0)$ . We follow the proofs of [11] and [20], but we take into consideration the special role of the point  $(0, 0)$ , which is a singular point for some of the  $q_j$ .

**Definition 6.1.** We say that a probability measure  $\nu$  on  $\mathbb{R}_+ \times \mathbb{R}_+$  belongs to  $\mathcal{P}_0$  when  $\nu(\{0, 0\}) = 0$ .

We will show that (i) for any sequence of initial distribution  $\nu^N$  converging weakly to  $\nu$  with  $\nu \in \mathcal{P}_0$ ,  $P_{Y^N, \nu^N}$  converges weakly to  $P_{\Psi, \nu}$ , (ii) for any probability measure  $\nu \in \mathcal{P}_0$ ,  $P_{\Psi(\cdot, t), \nu}$  converges weakly to  $\delta_{\psi^\infty}$ , (iii) the sequence  $(Y^N(\infty), N \geq 1)$  is tight, and that (iv) any possible accumulation point of  $(Y^N(\infty), N \geq 1)$  belongs to  $\mathcal{P}_0$ .

The proof is then short and elegant: since  $(Y^N(\infty), n \geq 1)$  is tight, it is sufficient to prove that there is a unique possible limit to any convergent subsequence of  $(Y^N(\infty))$ . We still denote by  $Y^N(\infty)$  such a converging subsequence (as  $N$  tends to  $\infty$ ). Its limit is denoted by  $\nu$ , known to belong to  $\mathcal{P}_0$ . According to (i) above,  $P_{Y^N, P_{Y^N(\infty)}}$  converges weakly to  $P_{\Psi, \nu}$ . Moreover, by the properties of Markov processes,  $P_{Y^N, P_{Y^N(\infty)}}$  is the distribution of a stationary process; hence,  $\Psi$  is also a stationary process when started from  $\nu$ . This means that the distribution of  $\Psi(\cdot, t)$  is  $\nu$  for any  $t$ . Then, by (ii) above,  $\nu = \delta_{\psi^\infty}$ . We have thus proved that any convergent subsequence of  $Y^N(\infty)$  converges to  $\delta_{\psi^\infty}$ . We now turn to the proofs of the three necessary theorems.

**Theorem 6.1.** For any sequence of initial distribution  $\nu^N$  converging weakly to  $\nu \in \mathcal{P}_0$ ,  $P_{Y^N, \nu^N}$  converges weakly to  $P_{\Psi, \nu}$ .

*Proof.* We will proceed in two steps: first we prove the tightness in  $\mathbb{D}([0, T], \mathbb{R}^2)$  and then identify the limit. In fact, we will prove the slightly stronger results that  $P_{Y^N, \nu^N}$  is tight and that the limiting process is continuous. According to [4], we need to show that, for each positive  $\epsilon$  and  $\eta$ , there exist  $\delta > 0$  and  $n_0$  such that, for any  $N \geq n_0$ ,

$$P\left(\sup_{|v-u| \leq \delta; v, u \leq T} \|Y^N(v) - Y^N(u)\| \geq \epsilon\right) \leq \eta.$$

We define

$$\begin{aligned}
 A_1^N(t) &= \frac{1}{N} \int_0^t (q_1^N + q_3^N - q_2^N - q_6^N)(X^N(s)) \, ds, \\
 A_2^N(t) &= \frac{1}{N} \int_0^t (q_4^N - q_3^N - q_5^N + q_6^N)(X^N(s)) \, ds.
 \end{aligned}$$

From Theorem 5.2 we know that

$$Y_i^N(v) = A_i^N(v) + \frac{1}{N} M_i^N(v), \quad i = 1, 2.$$

Hence, for any positive  $a$ ,

$$\begin{aligned}
 & \mathbb{P}\left(\sup_{|v-u|\leq\delta; v,u\leq T} \|Y^N(v) - Y^N(u)\| \geq \epsilon\right) \\
 & \leq \mathbb{P}(\|Y^N(0)\| \geq a) \\
 & \quad + \mathbb{P}\left(\sup_{|v-u|\leq\delta; v,u\leq T} \|A^N(v) - A^N(u)\| \geq \frac{\epsilon}{2}, \|Y^N(0)\| \leq a\right) \\
 & \quad + \mathbb{P}\left(\sup_{|v-u|\leq\delta; v,u\leq T} \frac{1}{N} \|M^N(v) - M^N(u)\| \geq \frac{\epsilon}{2}, \|Y^N(0)\| \leq a\right). \tag{6.1}
 \end{aligned}$$

Equation (5.2) implies that

$$\mathbb{E}\left[\sup_{s\leq T} \frac{1}{N^2} \|M^N(s)\|^2 \mid \|Y^N(0)\| \leq a\right] \leq \frac{C(a+1)}{N}.$$

This means that  $(N^{-1}M^N, N \geq 1)$  converges to 0 in  $L^2(\Omega; \mathbb{D}([0, T], \mathbb{R}^2), \mathbb{P}_{\cdot|\|Y^N(0)\|\leq a})$ . Hence, it converges in distribution in  $\mathbb{D}([0, T], \mathbb{R}^2)$  and, thus, is tight. This means that the last summand of (6.1) can be made as small as needed for large  $N$ . Furthermore,

$$\begin{aligned}
 \|A^N(v) - A^N(u)\| & \leq \frac{2}{N} \int_u^v \sum_{i=1}^6 q_i(X^N(s)) ds \\
 & \leq 2|v - u| \left( \frac{r_N + \lambda_N}{N} + \frac{C}{N} \sup_{s\leq T} \|X^N(s)\| \right).
 \end{aligned}$$

It follows from Lemma 3.1 that

$$\begin{aligned}
 \mathbb{E}\left[\sup_{|v-u|\leq\delta; v,u\leq T} \|A^N(v) - A^N(u)\| \mid \|Y^N(0)\| \leq a\right] & \leq C\delta \left( \frac{r_N + \lambda_N}{N} T + a \right) \\
 & \leq C((r + \lambda)T + a)\delta.
 \end{aligned}$$

This means that the second summand of (6.1) can also be made as small as needed. The hypothesis on the initial condition means that this also holds for the first summand of (6.1). Thus, we have proved so far that  $\mathbb{P}_{Y^N, v^N}$  is tight and that its limit belongs to the space of continuous functions.

We now prove that the only possible limit is  $\mathbb{P}_{\Psi, v}$ . Assume that  $v^N$  tends to  $v$  and that  $\mathbb{P}_{Y^N, v^N}$  tends to some  $\mathbb{P}_{Z, v}$ . We suppose that the initial conditions  $X^N(0)$  of the Markov processes are distributed as  $v_N$  and we introduce a random variable  $x^0$  distributed as  $v$ . Recall that  $Y^N = N^{-1}X^N$ . We fix  $M \in \mathbb{N}^*$ ,  $(\alpha^k = (\alpha_1^k, \alpha_2^k))_{0\leq k\leq M} \in \mathbb{R}^{2M+2}$ , and  $0 = t_0 \leq t_1 \leq \dots \leq t_M$ . We introduce

$$\begin{aligned}
 G^N & = \mathbb{E}\left[\exp i\left(\sum_{k=0}^M \langle \alpha_k, Y^N(t_k) \rangle\right)\right], \\
 \tilde{G}^N & = \mathbb{E}\left[\exp i\left(\sum_{k=0}^M \left\langle \alpha_k, \Psi\left(\frac{X^N(0)}{N}, t_k\right)\right\rangle\right)\right], \\
 G & = \mathbb{E}\left[\exp i\left(\sum_{k=0}^M \langle \alpha_k, \Psi(x^0, t_k) \rangle\right)\right],
 \end{aligned}$$

where  $Y^N(-1) = 0$  and  $Y^N = N^{-1}X^N$ , with initial condition  $X^N(0)$  distributed as  $\nu_N$  and  $X^0$  as  $\nu$ .

Let  $\varepsilon > 0$ . The sequence  $(\nu_N)_{N \in \mathbb{N}}$  is tight; hence, there exists a compact set  $K \subset \mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}$  such that  $\nu(K^c) + \sup_N \nu_N(K^c) \leq \varepsilon$ . We also introduce

$$G_K^N = \mathbb{E} \left[ \exp i \left( \sum_{k=0}^M \langle \alpha_k, Y^N(t_k) \rangle \right) \mathbf{1}_K \left( \frac{X^N(0)}{N} \right) \right],$$

$$\tilde{G}_K^N = \mathbb{E} \left[ \exp i \left( \sum_{k=0}^M \left\langle \alpha_k, \Psi \left( \frac{X^N(0)}{N}, t_k \right) \right\rangle \right) \mathbf{1}_K \left( \frac{X^N(0)}{N} \right) \right],$$

$$G_K = \mathbb{E} \left[ \exp i \left( \sum_{k=0}^M \langle \alpha_k, \Psi(x^0, t_k) \rangle \right) \mathbf{1}_K(x^0) \right].$$

Then,

$$\limsup_N |G - G^N| \leq 2\varepsilon + \limsup_N |\tilde{G}_K^N - G_K^N| + \limsup_N |\tilde{G}_K^N - G_K|.$$

From Theorem 4.1, the map  $(x, s) \mapsto \Psi(x, s)$  is continuous on  $(\mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}) \times [0, T]$  and  $\inf_{(x,s) \in K \times [0,T]} \|\Psi(x, s)\| > 0$ . Since  $X^N(0)/N$  takes values in the compact set  $K$ , then, from Lemma 5.1,  $\limsup_N |\tilde{G}_K^N - G_K^N| = 0$ . Since the sequence of measures  $(\nu_N)$  converges weakly to  $\nu$ , then  $\limsup_N |\tilde{G}_K^N - G_K| = 0$ . Hence,

$$\limsup_N |G - G^N| \leq 2\varepsilon \quad \text{for all } \varepsilon > 0.$$

This means that for any  $t_0, \dots, t_M$ ,

$$\mathbb{P}_{(Y^N(t_0), \dots, Y^N(t_M)), \nu_N} \rightarrow \mathbb{P}_{(\Psi(x^0, t_0), \dots, \Psi(x^0, t_M)), \nu} \quad \text{as } N \rightarrow \infty.$$

Hence, all the accumulation points are the same and the convergence of  $\mathbb{P}_{Y^N, \nu_N}$  follows.

**Theorem 6.2.** *For any probability measure  $\nu \in \mathcal{P}_0$ ,  $\mathbb{P}_{\Psi(\cdot, t), \nu}$  converges weakly to  $\delta_{\psi^\infty}$  as  $t \rightarrow \infty$ .*

*Proof.* For any  $f$  continuous bounded function on  $\mathbb{R}^2$ , we have

$$\int f \, d\mathbb{P}_{\Psi(\cdot, t), \nu} = \int_{\mathbb{R}^2} f(\Psi(x, t)) \, d\nu(x).$$

Theorem 4.1 says that, for any  $x \in \mathbb{R}_+ \times \mathbb{R}_+ \setminus \{(0, 0)\}$ ,

$$f(\Psi(x, t)) \rightarrow f(\psi^\infty) \quad \text{as } t \rightarrow \infty.$$

The result follows by dominated convergence.

**Theorem 6.3.** *The sequence  $(Y^N(\infty), N \geq 1)$  is tight and any accumulation point belongs to  $\mathcal{P}_0$ .*

We need a preliminary lemma which relies on the observation that when  $\mu_1 = \mu_2$ , the process  $X_1 + X_2$  has the dynamics of the process counting the number of customers in an

M/M/∞ queue. Recall that  $\mu_- = \min(\gamma\mu_1, \mu_2)$ , and set  $\zeta = (r + \lambda p_1)/\mu_-$ . For any  $c \in \mathbb{R}_+$  and any  $x \in \mathbb{N}$ , define the function

$$h_c(t, x) = (1 + ce^{\mu-t})^x e^{-\zeta c \exp(\mu-t)}.$$

Note that  $h_c$  is increasing with respect to  $x$ . Moreover, according to [17, Chapter 6],

$$\frac{\partial h_c}{\partial t}(t, x) + R(h_c(t, \cdot))(x) = 0, \tag{6.2}$$

where, for any  $w : \mathbb{N} \rightarrow \mathbb{R}$ ,

$$Rw(x) = (w(x + 1) - w(x))(r + \lambda p_1) + (w(x - 1) - w(x))\mu_- x.$$

**Lemma 6.1.** *For any nonnegative real  $c$ , the process  $H_c = (h_c(t, X_1(t) + X_2(t)), t \geq 0)$  is a positive supermartingale.*

*Proof.* According to Dynkin’s formula (see [17, Proposition C.5]), for any  $0 \leq s < t$ , we have

$$\begin{aligned} 0 &= \mathbb{E} \left[ h_c(t, \|X(t)\|) - h_c(s, \|X(s)\|) - \int_s^t \frac{\partial h_c}{\partial t}(r, \|X(r)\|) \, dr \right. \\ &\quad - (r + \lambda p_1) \int_s^t (h_c(r, \|X(r)\| + 1) - h_c(r, \|X(r)\|)) \, dr \\ &\quad \left. - \int_s^t (h_c(r, \|X(r)\| - 1) - h_c(r, \|X(r)\|))(\gamma\mu_1 X_1(r) + \mu_2 X_2(r)) \, dr \mid \mathcal{F}_s \right] \\ &\geq \mathbb{E} \left[ h_c(t, \|X(t)\|) - h_c(s, \|X(s)\|) - \int_s^t \frac{\partial h_c}{\partial t}(r, \|X(r)\|) \, dr \right. \\ &\quad - (r + \lambda p_1) \int_s^t (h_c(r, \|X(r)\| + 1) - h_c(r, \|X(r)\|)) \, dr \\ &\quad \left. - \int_s^t (h_c(r, \|X(r)\| - 1) - h_c(r, \|X(r)\|))\mu_-(X_1(r) + X_2(r)) \, dr \mid \mathcal{F}_s \right], \end{aligned}$$

where the inequality follows from the monotonicity of  $h_c$  and the definition of  $\mu_-$ . Hence, we obtain

$$\begin{aligned} 0 &\geq \mathbb{E} \left[ h_c(t, \|X(t)\|) - h_c(s, \|X(s)\|) \right. \\ &\quad \left. - \int_s^t \left( \frac{\partial h_c}{\partial t}(r, \|X(r)\|) + R(h_c(r, \cdot))(\|X(r)\|) \right) \, dr \mid \mathcal{F}_s \right]. \end{aligned}$$

In view of (6.2), we obtain

$$0 \geq \mathbb{E}[h_c(t, \|X(t)\|) - h_c(s, \|X(s)\|) \mid \mathcal{F}_s],$$

i.e.  $H_c$  is a supermartingale.

Now, let  $(Y^N(\infty), N \geq 1)$  be a subsequence which converges to  $\nu$ . Since  $X^N(\infty)$  is a random variable distributed according to the stationary law of the process  $X^N$ ,

$$\mathbb{E}[Q^N e^{-\|\cdot\|}(X^N(\infty))] = 0,$$

where  $Q^N$  is the infinitesimal generator of  $X^N$ . By a direct calculation we have

$$Q^N e^{-\|\cdot\|}(x) = e^{-\|x\|}[(\lambda_N + r_N)(e^{-1} - 1) + (\gamma\mu_1x_1 + \mu_2x_2)(e - 1)].$$

Then

$$\begin{aligned} &(\lambda_N + r_N)(1 - e^{-1}) E[e^{-N\|Y^N(\infty)\|}] \\ &= N(e - 1) E[e^{-N\|Y^N(\infty)\|}(\gamma\mu_1Y_1^N(\infty) + \mu_2Y_2^N(\infty))]. \end{aligned}$$

Hence,  $(1 - e^{-1})(r + \lambda)v(\{(0, 0)\}) = 0$ , i.e.  $v$  belongs to  $\mathcal{P}_0$ . This completes the proof.

*Proof of Theorem 6.3.* Let  $K$  be real. For any positive real  $\theta$ , we have

$$P(\|Y^N(t)\| > K) = P(\|X^N(t)\| > NK) \leq e^{-\theta NK} E[\exp(\theta\|X^N(t)\|)].$$

Lemma 6.1 entails that

$$E[\exp(\theta\|X^N(t)\|)] \leq (1 + (e^\theta - 1)e^{-\mu-t})^{N\|X^N(0)\|} \exp(N\zeta(e^\theta - 1)(1 - e^{-\mu-t})).$$

Hence,

$$\begin{aligned} P(\|Y^N(\infty)\| > K) &= \lim_{t \rightarrow \infty} P(\|Y^N(t)\| > K) \\ &\leq \inf_{\theta > 0} \lim_{t \rightarrow \infty} (1 + (e^\theta - 1)e^{-\mu-t})^{N\|X^N(0)\|} \\ &\quad \times \exp(-\theta NK + N\zeta(e^\theta - 1)(1 - e^{-\mu-t})) \\ &= \inf_{\theta > 0} \exp(N(-\theta K + \zeta(e^\theta - 1))) \\ &\leq \exp(-\frac{1}{2}NK \ln K) \end{aligned}$$

for large enough  $K$ . The tightness follows.

### 7. Central limit theorem

It turns out that we can also evaluate the order of the approximation when we replace  $X^N$  by  $\Psi$ . This is given by a central limit theorem such as the following.

**Theorem 7.1.** *Assume that the hypothesis of Theorem 5.1 holds. Assume further that  $\sqrt{N}(Y^N(0) - x^0)$  tends to 0 as  $N$  tends to  $\infty$ . Then, for any  $T > 0$ , the process*

$$W^N = \sqrt{N}(Y^N - \Psi(x^0, \cdot))$$

*tends in distribution in  $\mathbb{D}([0, T], \mathbb{R}^2)$  to a centered Gaussian process with covariance matrix  $\Gamma(t)$  given by*

$$\Gamma(t) = \begin{pmatrix} \Gamma_1(t) & \Gamma_{12}(t) \\ \Gamma_{12}(t) & \Gamma_2(t) \end{pmatrix},$$

where

$$\begin{aligned} \Gamma_1(t) &= rt + \int_0^t \left( \lambda p_1 \frac{\Psi_1(x^0, s)}{\Psi_1(x^0, s) + \Psi_2(x^0, s)} + \mu_1 \Psi_1(x^0, s) \right. \\ &\quad \left. + \alpha p \frac{\Psi_1(x^0, s)\Psi_2(x^0, s)}{\Psi_1(x^0, s) + \Psi_2(x^0, s)} \right) ds, \\ \Gamma_2(t) &= \int_0^t \left( \lambda \left( 1 - p_1 \frac{\Psi_1(x^0, s)}{\Psi_1(x^0, s) + \Psi_2(x^0, s)} \right) + \mu_2 \Psi_2(x^0, s) \right. \\ &\quad \left. + \alpha p \frac{\Psi_1(x^0, s)\Psi_2(x^0, s)}{\Psi_1(x^0, s) + \Psi_2(x^0, s)} \right) ds, \\ \Gamma_{12}(t) &= -\alpha p \int_0^t \frac{\Psi_1(x^0, s)\Psi_2(x^0, s)}{\Psi_1(x^0, s) + \Psi_2(x^0, s)} ds - (1 - \gamma)\mu_1 \int_0^t \Psi_1(x^0, s) ds. \end{aligned}$$

*Proof.* According to [9, p. 339], it suffices to prove that

$$E \left[ \sup_{t \leq T} |W^N(t) - W^N(t_-)| \right] \rightarrow 0 \quad \text{as } N \rightarrow +\infty$$

and that

$$\langle\langle W^N \rangle\rangle_t \rightarrow \Gamma(t) \quad \text{as } N \rightarrow +\infty.$$

Since the jumps of  $Y^N$  are bounded by  $1/N$ , those of  $W^N$  are bounded by  $N^{-1/2}$ ; hence, the first point is proved. As to the second point, note that

$$\langle\langle W^N \rangle\rangle_t = N^{-1} \langle\langle M^N \rangle\rangle_t$$

and then use Theorem 5.1.

### 8. Numerical investigation

Another approach to evaluate the order of approximation can be made by computer simulation. We simulate the Markov process for  $N = 100$  and compute the estimate of the prevalence by a simple Monte Carlo method on 10 000 trajectories. For each sample path, the prevalence is evaluated as the prevalence at a time large enough to have presumably attained the stationary regime, i.e. we chose to stop after 10 000 transitions. For the parameters we chose,  $\alpha = 1$ ,  $\mu_1 = 0.1$ ,  $\mu_2 = 0.2$ ,  $r = 1$ ,  $\lambda = 5$ , and  $p_1 = 0.8$ , the results are strikingly good, as shown in Figure 3. Note that the choice of parameters here is very delicate, since biological parameters are not very well known (i.e.  $\mu_1, \mu_2, p$ , etc.) and population-dependent quantities are even more obscure to determine. Here we chose parameters which seem reasonable and fit the observed prevalence. Note that the initial condition does not matter since we are looking at the long-term behavior. Since we have a Markov process, it is well known that the starting point does not influence the stationary regime.

In such models, another quantity of interest is the relative importance of each parameter, i.e. what affects the prevalence the most? In the deterministic system, this question is easily solved by computing the derivative of the prevalence with respect to each of the parameters. We now explain how to compute the sensitivity of the prevalence in the stochastic model. Say we have a bounded function  $F$  which depends on the sample paths of  $X$ . We aim to compute

$$\frac{d}{dp} E_p[F],$$

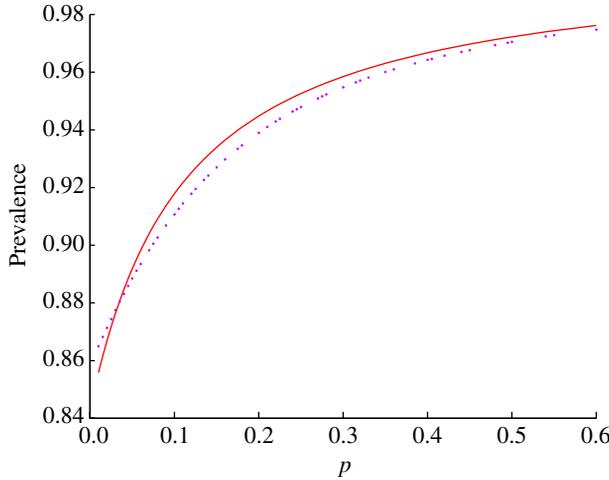


FIGURE 3: Prevalence with respect to  $p$ . The solid line represents the value as computed by (4.1). The dots represent the simulated values. The 95% confidence intervals are so small that they cannot be displayed.

where we put a  $p$  under the expectation symbol to emphasize the dependence of the underlying probability with respect to  $p$ . Other ‘greeks’, as these quantities are called in mathematical finance, can be derived analogously. We assume that we observe the Markov process on a time window of size  $T$ , i.e. any functional is implicitly assumed to belong to  $\mathcal{F}_T = \sigma\{X(s), 0 \leq s \leq T\}$ .

**Theorem 8.1.** *For any bounded  $F$ ,  $F \in \mathcal{F}_T$ , we have*

$$\begin{aligned} \frac{d}{dp} E_p[F] &= \frac{1}{p} E_p \left[ F \left( \sum_{s \leq T} \mathbf{1}_{\{(1,-1)\}}(\Delta X(s)) - \int_0^T q_3(X(s_-)) ds \right) \right] \\ &= \frac{1}{p} \text{cov}_p \left( F, \sum_{s \leq T} \mathbf{1}_{\{(1,-1)\}}(\Delta X(s)) \right), \end{aligned}$$

where  $\Delta X(s) = X(s) - X(s_-)$ .

*Proof.* The proof relies on the Girsanov theorem, which is more easily expressed in the framework of multivariate point measures. Since there are only six kinds of jump, we can represent the dynamics of  $X$  as a point measure on  $\mathbb{R}^+ \times \{1, \dots, 6\}$ :

$$\mu([0, t] \times \{i\}) = \sum_{s \leq t} \mathbf{1}_{\{\Delta X(s)=l_i\}},$$

where

$$\begin{aligned} l_1 &= (1, 0), & l_2 &= (-1, 0), & l_3 &= (1, -1), \\ l_4 &= (0, 1), & l_5 &= (0, -1), & l_6 &= (-1, 1). \end{aligned}$$

In the reverse direction,

$$X(t) = X(0) + \sum_{i=1}^6 \mu([0, t] \times \{i\})l_i.$$

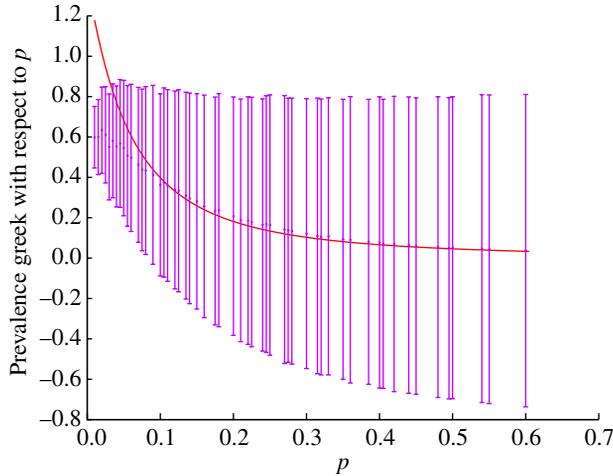


FIGURE 4: Prevalence greek with respect to  $p$ . Same conventions as above. The error bars represent the 95% confidence intervals.

It is immediate from the preceding results that  $v^p$ , the  $\mathbb{P}_p$ -predictable compensator of  $\mu$ , is given by  $dv^p(t, i) = q_i(X(t_-)) dt$ . To compute  $(d/dp)E_p[F]$  means to compute

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (E_{p+\varepsilon}[F] - E_p[F]).$$

Under  $\mathbb{P}_{p+\varepsilon}$ ,  $dv^{p+\varepsilon}(t, i) = dv^p(t, i)$  for  $i \neq 3$  and  $dv^{p+\varepsilon}(t, 3) = (1 + \varepsilon/p) dv^p(t, 3)$ . Let

$$U(t, i) = \begin{cases} 0 & \text{if } i \neq 3, \\ \varepsilon/p & \text{if } i = 3. \end{cases}$$

According to the Girsanov theorem (see [7] and [14]), this means that

$$\begin{aligned} E_{p+\varepsilon}[F] &= E_p \left[ F \mathcal{E} \left( \sum_{i=1}^6 \int_0^t U(s, i) (d\mu(s, i) - dv^p(s, i)) \right) \right] \\ &= E_p \left[ F \mathcal{E} \left( \frac{\varepsilon}{p} (\mu([0, T] \times \{3\}) - v^p([0, T] \times \{3\})) \right) \right], \end{aligned}$$

where  $\mathcal{E}$  denotes the Doléans-Dade exponential. It is known (see [7]) that a Doléans-Dade exponential follows the same rule of derivation as a usual exponential. Moreover, since  $F$  is bounded, we can differentiate inside the expectation, completing the proof.

With the parameters above, the simulated greek coincides pretty well with the sensitivity computed by differentiating the expression of the stationary prevalence in the deterministic system; see Figure 4. However, as usual with this method and unlike the above simulations, the confidence intervals are rather large.

### 9. Conclusion

We analyzed the sensitivity with respect to all parameters  $p$  for different sets of values. For the tested values, it appears that the most important parameter is  $r$ , the rate of new infectious IDUs

coming from the outside. This means that a sort of quarantine would guarantee, in the long term, a decrease of the prevalence. However, this seems practically infeasible. More strikingly is the fact that the second sensitive parameter is  $\lambda$ , the rate of new IDUs. This means that an efficient way to diminish the prevalence would be to augment the number of IDUs and, thus, mathematically speaking, to increase the denominator in the prevalence expression, more than its numerator. Again, this solution is not socially acceptable. The third sensitive parameter is  $p$ , the probability of infection at each injection. The influence of  $\gamma$  seems negligible: for the whole range of possible values of  $\gamma$  from 0 to 1, the prevalence varies by less than 10%. Hence, the possible healing of hepatitis C, which would intuitively help to decrease the prevalence, is negligible. Our analysis thus corroborates the intuition of the epidemiologists, that the main difference between the two viruses is explained by their infection strength.

### Acknowledgements

This work was carried out during a stay of Laurent Decreusefond at Université Paris Descartes. He would like to thank Université Paris Descartes for warm hospitality.

We would like to thank Y. Le Strat and J. C. Thalabard for suggesting the problem and for helpful discussions. We would also like to thank an anonymous referee for his/her helpful comments to improve the presentation of this work.

### References

- [1] ANDERSSON, H. AND BRITTON, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*. Springer, New York.
- [2] ANDERSSON, P. AND LINDENSTRAND, D. (2010). A stochastic SIS infection model incorporating indirect transmission epidemic with demography: initial stages and time to extinction. Submitted.
- [3] BELL, J. *et al.* (1990). Hepatitis C virus in intravenous drug users. *Med. J. Austral.* **153**, 274–276.
- [4] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
- [5] CLANCY, D. (2005). A stochastic SIS infection model incorporating indirect transmission. *J. Appl. Prob.* **42**, 726–737.
- [6] CLANCY, D. AND POLLETT, P. K. (2003). A note on quasi-stationary distributions of birth–death processes and the SIS logistic epidemic. *J. Appl. Prob.* **40**, 821–825.
- [7] DECREUSEFOND, L. (1998). Perturbation analysis and Malliavin calculus. *Ann. Appl. Prob.* **8**, 496–523.
- [8] ESPOSITO, N. AND ROSSI, C. (2004). A nested-epidemic model for the spread of hepatitis C among injecting drug users. *Math. Biosci.* **188**, 29–45.
- [9] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.
- [10] GARFEIN, R. S. *et al.* (1996). Viral infections in short-term injection drug users: the prevalence of the hepatitis C, hepatitis B, human immunodeficiency, and human T-lymphotropic viruses. *Amer. J. Public Health* **86**, 655–661.
- [11] GRAHAM, C. (2000). Kinetic limits for large communication networks. In *Modeling in Applied Sciences*, Birkhäuser, Boston, MA, pp. 317–370.
- [12] HEALTH PROTECTION AGENCY (2004). Shooting up: infection among injecting drug users in the United Kingdom 2003. Tech. Rep., London.
- [13] JACOB, C. AND VIET, A. F. (2003). Epidemiological modeling in a branching population. Particular case of a general SIS model with two age classes. *Math. Biosci.* **182**, 93–111.
- [14] JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. Springer, Berlin.
- [15] JAUFFRET-ROUSTIDE, M. *et al.* (2006). Impact of a harm-reduction policy on HIV and hepatitis C virus transmission among drug users: recent French data—the ANRS-Coquelicot study. *Subst. Use Misuse* **41**, 1603–1621.
- [16] MÉTIVIER, M. (1988). *Stochastic Partial Differential Equations in Infinite-Dimensional Spaces*. Scuola Normale Superiore, Pisa.
- [17] ROBERT, P. (2003). *Stochastic Networks and Queues* (Appl. Math. **52**). Springer, Berlin.
- [18] VERHULST, F. (1996). *Nonlinear Differential Equations and Dynamical Systems*, 2nd edn. Springer, Berlin.
- [19] VICKERMAN, P., HICKMAN, M. AND JUDD, A. (2007). Modelling the impact on Hepatitis C transmission of reducing syringe sharing: London case study. *Internat. J. Epidemiology* **36**, 396–405.
- [20] WHITT, W. (1985). Blocking when service is required from several facilities simultaneously. *Bell Syst. Tech. J.* **64**, 1807–1856.